IN THE UNITED STATES PATENT AND TRADEMARK OFFICE (USPTO)	
Application Serial Number	10/601,741
Confirmation Number	9003
Filing Date	06/23/2003
Title of Application	Advanced Spam Detection Techniques
First Named Inventor	Bryan T Starbuck
Assignee	Microsoft Corporation
Group Art Unit	2444
Examiner	Djenane M Bayard
Attorney Docket Number	MS1-4099US

To: Commissioner for Patents
PO Box 1450
Alexandria, VA 22313-1450

From: Kayla D. Brant (Tel. 509-324-9256; Fax 509-323-8979)
Lee & Hayes, PLLC
601 W Riverside Ave, Suite 1400
Spokane, WA 99201

Customer Number 22801

Paper Correcting Appeal Brief

(Response to the Communication Re: Appeal Dated 04/07/09)

[0001] Applicant does not believe any additional fees are due at this time. However, Applicant hereby authorizes the Commissioner to charge any deficiency of fees and credit any overpayments to Deposit Account Number 12-0769.

[0002] Applicant submits herein a corrected Appeal Brief, Claims Appendix, which includes a corrected listing of the claims, as amended in the Amendment After Final dated February 22, 2008. Applicant also submits herein a corrected Appeal Brief, Status of Amendments to reflect that the Amendment After Final dated February 22, 2008 has been entered.

IV. Status of Amendments (37 C.F.R. §41.37(c)(1)(iv))

The amendments that were submitted in the Reply to Final Office Action dated January 10, 2008 have been entered. No further amendments have been submitted.

VIII. Claims Appendix (37 C.F.R. §41.37(c)(1)(viii))

1. A computer-implemented spam detection system comprising:

a message parsing component that identifies features relating to at least a portion of origination information of a message; and

a feature pairing component that combines the features into useful pairs, the features of the pairs are evaluated for consistency with respect to one another to determine if the message is spam.

2. The system of claim 1, each pair comprises at least one of the following:

at least one of a domain name and a host name in a MAIL FROM command;

at least one of a domain name and a host name in a HELO COMMAND;

at least one of an IP address and a subnet in a Received from header;

at least one of a domain name and a host name in a Display name;

at least one of a domain name and a host name in a Message From line; and

at least one time zone in a last Received from header.

3. The system of claim 2, the domain name is derived from the host name.

4. The system of claim 2, the subnet comprises one or more IP addresses that share a first number of bits in common.

5. The system of claim 1, a useful pair is anyone of a domain name and a host name from a Message From and from a HELO command.

6. The system of claim 1, a useful pair is a Display name domain name and host name and a Message From domain name and host name.

7. The system of claim 1, a useful pair is anyone of a domain name and a host name in a Message From and anyone of a Received from IP address and subnet.

8. The system of claim 1, a useful pair is a sender's alleged time zone and a Message From domain name.

9. The system of claim 1, a useful pair comprises a sender's type of mailing software and any one of a domain name, host name and user name derived from one of an SMTP command and a message header.

10. The system of claim 1, origination information comprises SMTP commands, the SMTP commands comprise a HELO command, a MAIL FROM command, and a DATA command.

11. The system of claim 10, the DATA command comprises a Message From line, sender's alleged time zone, and sender's mailing software.

12. The system of claim 1, further comprising a component that applies one or more heuristics consistently to mail messages to obtain consistent feature pairing.

42. A computer-implemented method that facilitates generating features for use in spam detection comprising:

receiving at least one message;

parsing at least a portion of a message to generate one or more features;

combining at least two features into pairs, each pair of features creates at least one additional feature, the features of each pair coinciding with one another;

using the pairs of features to train a machine learning spam filter regarding acceptable or unacceptable pairs; and

detecting a spam e-mail based at least in part on comparing one or more pairs of features in the e-mail to at least one pair in the machine learning spam filter.

43. The method of claim 42, the at least a portion of the message being parsed corresponds to origination information of the message.

44. The method of claim 42, each pair comprises at least one of the following:

at least one of a domain name and a host name in a MAIL FROM command;

at least one of a domain name and a host name in a HELO COMMAND;

at least one of an IP address and a subnet in a Received from header;

at least one of a domain name and a host name in a Display name;

at least one of a domain name and a host name in a Message From line; and

at least one time zone in a last Received from header.

45. The method of claim 44, the domain name is derived from the host name.

46. The method of claim 42, the pair of features is a Display name domain name and host name and a Message From domain name and host name.

47. The method of claim 42, a useful pair is anyone of a domain name and a host name from a Message From and from a HELO command.

48. The method of claim 42, the pair of features is anyone of a domain name and a host name in a Message From and anyone of a Received from IP address and subnet.

49. The method of claim 42, the pair of features is a sender's alleged time zone and a Message From domain name.

50. The method of claim 42, the pair of features comprises a sender's type of mailing software and anyone of a domain name, host name and display name derived from one of an SMTP command and a message header.

51. The method of claim 42, further comprising selecting one or more most useful pairs of features to train the machine learning filter.

52. The method of claim 42, the detecting a spam e-mail based at least in part on one of:

receiving new messages;

generating pairs of features based on origination information in the messages;

passing the pairs of features through the machine learning filter; and

obtaining a verdict as to whether at least one pair of features indicates that the message is more likely to be spam.

73. A computer-implemented system that facilitates generating features for use in spam detection comprising:

means for receiving at least one message;

means for parsing at least a portion of a message to generate one or more features;

means for combining at least two features into pairs, the pairs are evaluated against each other for consistency; and

means for using the pairs of features to train a machine learning spam filter.

Conclusion

[0003] Please contact the undersigned representative for the Applicant if any further issues remain.

Respectfully Submitted,

Lee & Hayes, PLLC
Representative for Applicant

/Kayla D. Brant #46,576/ Dated: June 2, 2009

Kayla D. Brant
(kayla@leehayes.com; 509-944-4742)
Registration No. 46576